ELSEVIER

Aquatic Botany 62 (1999) 269–283

Aquatic botany

Chloroplast tRNA^{Leu} (UAA) intron sequences provide phylogenetic resolution of seagrass relationships

Gabriele Procaccini[a,*], Lucia Mazzella[a], Randall S. Alberte[b], Donald H. Les[c]

[a]*Laboratorio di Ecologia del Benthos, Stazione Zoologica 'A. Dohrn'; 80077, Ischia (Napoli), Italy*
[b]*Aretè Associates, Arlington VA, USA*
[c]*Department of Ecology and Evolutionary Biology, The University of Connecticut, Storrs CT06269-3043, USA*

Received 1 December 1997; accepted 11 August 1998

Abstract

Recent molecular studies have indicated that seagrasses comprise three convergent angiosperm clades. Although seagrass polyphyly has been demonstrated persuasively, other details of their phylogenetic relationships remain uncertain or weakly supported. To further assess seagrass relationships, we explored the potential of chloroplast *trn*L (UAA) intron sequences for phylogenetic reconstruction in the Alismatidae. Sequence analysis revealed considerable length variation of the *trn*L intron among the eight species of the subclass Alismatidae examined. These regions (representing large insertions/deletions in loops) were difficult to align and too variable to use reliably in phylogenetic analysis. However, conserved regions of the intron were readily aligned and were characterized by levels of divergence comparable to coding *rbc*L sequences. When analyzed phylogenetically, conserved *trn*L intron sequences recovered the same phylogenetic relationships among seagrass clades that were obtained using *rbc*L data. Combined analysis of *trn*L intron and *rbc*L coding sequences yielded a single most parsimonious tree with levels of nodal support higher than those obtained independently for either of the datasets. Analyses of conserved intron and coding chloroplast DNA sequences provide continued support for the polyphyly of seagrasses, the monophyly of Zosteraceae and a clade comprising Ruppiaceae, Posidoniaceae and Cymodoceaceae. Conserved *trn*L intron should be useful for evaluating other phylogenetic relationships in subclass Alismatidae.

Keywords: Seagrass; Chloroplast DNA; *Trn*L intron; *Rbc*L; Phylogenetics; Parsimony maximum likelihood

* Corresponding author. Tel.: +39 081 5833 514/5833259; fax: +39 081 984201; e-mail: gpro@alpha.szn.it

0304-3770/99/$ – see front matter
PII: S0304-3770(98)00099-0

1. Introduction

Seagrasses (marine angiosperms) form extensive beds in shallow waters of all temperate and tropical seas (den Hartog, 1970; Larkum and den Hartog, 1989). Based on non-molecular characters, these marine angiosperms have been classified within the monocotyledon subclass Alismatidae (Cronquist, 1988; Kuo and McComb, 1989). Molecular phylogenetic analyses using gene sequences of the ribulose bisphosphate carboxylase large subunit (*rbc*L) support the inclusion of seagrasses in the subclass Alismatidae and further indicate that they comprise three convergent clades with fresh or brackish water sister-groups (Les and Haynes, 1995; Les et al., 1993, 1997; Waycott and Les, 1996). Although seagrass polyphyly has been forcefully indicated by both morphological and molecular datasets, various details of their phylogenetic relationships remain unsettled or inadequately supported.

Chloroplast DNA (cpDNA) variation has been used extensively to investigate phyletic relationships among plants (Giannasi et al., 1992; Chase et al., 1993; Yong-Pyo et al., 1993), hybrid zones and long distance gene flow (Soltis et al., 1992). Previous studies based on restriction analysis and/or gene sequences suggest that chloroplast genes have slow rates of nucleotide substitution with an average synonymous rate almost half that of plant nuclear DNA (Wolfe et al., 1987; Zurawski and Clegg, 1987). However, non-coding regions of cpDNA exhibit faster mutation rates, especially through the addition and/or deletion of short sequences (Zurawski and Clegg, 1987). Conceivably, such regions may potentially complement more slowly evolving coding cpDNA sequences by providing a greater number of genetic markers, especially at the level of infrageneric comparisons.

One region shown to be particularly useful at detecting species level variation is the *trn*L (UAA) cpDNA intron (Taberlet et al., 1991; Ferris et al., 1993; Gielly and Taberlet, 1994). In this paper, we explore the potential use of this intron for resolving phyletic relationships among eight species of subclass Alismatidae, including members of all five extant seagrass families.

2. Materials and methods

2.1. Plant collection and DNA extraction

Tissue samples and DNA were obtained from eight species belonging to the subclass Alismatidae (Table 1). DNA was extracted with 2 × CTAB buffer, as in Procaccini et al. (1996). The method yielded 50–100 $\mu g\ g^{-1}$ fresh tissue of DNA for all species examined.

2.2. PCR amplification of the trnL (UAA) chloroplast intron

PCR primers were synthesized (Applied BioSystems 480A) based on sequences described by Taberlet et al. (1991). PCR amplifications were performed in a total volume of 100 µl (final concentrations: 50 mM KCl, 10 mM Tris–HCl pH 9.0, 1% (w/v) Triton X-100, 0.4 mM of each nucleotide, 5.0 mM $MgCl_2$, 1 µM of the appropriate primers,

Table 1
Species of the subclass Alismatidae examined and collection sites

Family	Genus	Species	Collection site
Cymodoceaceae	*Cymodocea*[a]	*C. nodosa* (Ucria) Asch.	Island of Ischia, Gulf of Naples (Italy)
Hydrocharitaceae	*Halophila*[a]	*H. stipulacea* (Forssk.) Asch.	Taormina, Sicily (Italy)
Juncaginaceae	*Triglochin*[c]	*T. maritimum* L.	Wisconsin (USA)
Posidoniaceae	*Posidonia*[a]	*P. oceanica* (L.) Delile	Island of Ischia, Gulf of Naples (Italy)
Ruppiaceae	*Ruppia*[b]	*R. maritima* L.	Monterey Bay (Calfornia, USA)
Zosteraceae	*Phyllospadix*[a]	*P. torreyi* Hook.	Monterey Bay (Calfornia, USA)
	Zostera[a]	*Z. marina* L.	Monterey Bay (California, USA)
		Z. noltii Horn.	Island of Ischia, Gulf of Naples (Italy)

Family designations as in Cronquist (1988).
[a] Marine genus.
[b] Brackish water genus.
[c] Fresh-water genus.

2.5 U Taq Polymerase) in the presence of 50–100 ng of template. Amplifications were carried out using a 'hot start' (94°C for 15 min), followed by 35 cycles (1.5 min 94°C, 1 min 55°C, 3 min 72°C) and 7 min extension at 72°C.

2.3. DNA sequencing

PCR products were separated electrophoretically in 1% agarose gels, excised from the gel, and purified by centrifugation through siliconized glass wool (Heery et al., 1990) or by GENECLEAN (Bio 101) in accordance with the manufacturer's instructions. Single-stranded DNA was sequenced in both directions using either the Delta Taq Cycle Sequencing Kit (USB) or the Thermo Sequenase Cycle Sequencing Kit (Amersham). The complete sequence of the group I intron was obtained by using the primers indicated in Table 2.

Table 2
Primer sequences and locations

Name	Sequence 5′–3′	Position
C	CGAAATCGGTAGACGCTACG	External
C1	GCAGAGACTCAATGGAAGCT	316
C2	CGTACATATACATACTGACA	443
D	GGGGATAGAGGGACTTGAAC	External
D1	TGTCAGTATGTATATGTACG	462
D2	AGCTTCCATTGAGTCTCTGC	335
D3	TAGGATAGGGTCATCCTTCC	426
D4	TATAAGGTCATTCTTCCTTTCGGC	427

Primers C and D (Taberlet et al., 1991) were used for amplification.
'C' primers read in 5′–3′ direction and 'D' primers read in 3′–5′ direction. The position refers to the first base of the primer in the sequence alignment in Fig. 1. Primers D3 and D4 are specific for *P. oceanica* and *H. stipulacea*, respectively.

2.4. Phylogenetic analyses

Preliminary multiple alignments of the sequences of the *trn*L (UAA) intron were obtained using CLUSTAL (Higgins et al., 1992) with manual correction for large gaps. The CLUSTAL alignment was checked and refined from independent examination by two of the authors. Ambiguous sequences (regions that could not be confidently aligned across all taxa) were eliminated from the final *trn*L data set which included 472 relatively conserved positions (Fig. 1). Gaps were treated as missing characters. A second dataset comprising previously published *rbc*L sequences (1183 nucleotides; GenBank accession numbers: U03729, U80688, U80698, U80714, U80719, U80731, U80733, U80734) was analyzed separately and in combination with the *trn*L data.

Phylogenetic analysis of the eight taxa was conducted using PAUP* test version 4.0.0d57 (Swofford, 1993; use of test version by permission). Pairwise distances (uncorrected '*p*' distance) were calculated between all taxa for *trn*L intron sequences, *rbc*L coding sequences and the combined *trn*L/*rbc*L data. The consistency index (CI), the consistency index in exclusion of uninformative sites ($CI_{(EXC)}$), retention index (RI) and g_1 (skewness) values were calculated for all parsimony analyses (g_1 was calculated directly using exhaustive searches).

Several analyses were performed. *Halophila* was used as an outgroup (in accordance with Les et al., 1997) in unweighted and weighted parsimony analysis where a stepmatrix was used to downweight transitions in a 1 : 2 ratio to transversions. To reduce effects of long branch attraction, unweighted parsimony analysis was repeated in exclusion of the divergent *Halophila* sequence with *Triglochin* redefined as the outgroup (as indicated by the weighted parsimony analysis). All searches were exhaustive. Bootstrap support was obtained from 500 replicates using branch and bound search. Strict consensus was used to depict results where several equally parsimonious trees resulted.

The *trn*L data were evaluated further using maximum likelihood estimation (MLE). Two models were examined: F81 (all sites assumed to evolve at the same rate) and HKY (two substitution types with transition/transversion ratio and proportion of invariant sites estimated by maximum likelihood). Initial branch lengths were obtained by Rogers–Swofford approximation. MLE searches used heuristic search with TBR branch swapping. Bootstrap support was obtained as described above.

Previously published *rbc*L sequences were reanalyzed for this set of eight taxa using the weighting scheme followed by Les et al. (1997) which adjusts for transition bias and codon positions (search procedures followed parsimony analyses described above). Congruency of *trn*L and *rbc*L data was evaluated by calculating the Mickevich–Farris index of incongruency (I_{MF}) following the methods outlined in Swofford (1991). The *rbc*L and *trn*L data were combined and analyzed using weighted and unweighted parsimony and maximum likelihood analyses incorporating F81 and HKY models. Bootstrap values were obtained as described above.

3. Results

The size of the $tRNA^{Leu}$ (UAA) intron was determined by starting from putative splicing sites (as in Kuhsel et al., 1990) and ranged from 574 bp in *Triglochin maritimum* to 723 bp in *Halophila stipulacea* (Fig. 1).

```
             [        10        20        30        40      *****P******60        70         ****Q*****     ]
C.n.  [gactt] GATTGGATTGAGCCTTAGTATGGAAACCTACTAAGTGGTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAAGGGGCAATCCTGAGCCAAAT
H.s.  [gactt] GATTGGATTGAGCCTTGGTATGGAAACCTACTAAGTGATAACTTTCAAATTCAGAGAAACCCTGGAATGAAAAATGGGCAATCCTGAGCCAAAT
P.o.  [gactt] AATTGGATTGAGCCTTGGTATGGAAACCTACTAAGTGGTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAATGGGCAATCCTGAGCCAAAT
P.t.  [gactt] GATTGGATTGAGCCTTGGTATGGAAACCTACTAAGTGTTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAATGGGCAATCCTGAGCCAAAT
R.m.  [gactt] GATTGGATTGAGCCTTGGTAGGGAAACCTACTAAGTGGTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAATGGGCAATCCTGAGCCAAAT
T.m.  [gactt] GATTGGATTGAGCCTTGGTATGGAAACCTACTAAGTGGTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAATGGGCAATCCTGAGCCAAAT
Z.m.  [gactt] AATTGGATTGAGCATTGGTATGGAAACCTACTAAGTGTTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAATGGGCAATCCTGAGCCAAAT
Z.n.  [gactt] AATTGGATTGAGCATTGGTATGGAAACCTACTAAGTGTTAACTTCCAAATTCAGAGAAACCCTGGAATTAAAAATGGGCAATCCTGAGCCAAAT

[         100       110         120       130       140       150       160       170       180       190   ]
C.n.  CCTTGTTTTGACAAAAAA [AGTTTTATTTTCTATATATAGAATTATATATAGAATATAGAATTTTATATAGAAATTATAAATTC-----------------
H.s.  CCTTGTTTTGAAAAAAAA [AAAAAAAAAAAAAGTTTCTTATCTTAGAACAAGGTACATAAATAGATAGGAAATCTATATGCATATTTCCATCTATTCTCTAT
P.o.  CCTTGTTTTGAGAAAAAA [GGGTTTATTTTCTATATTTATAATTTTATTTCTATTTTAGATATTTATATATATATATTATATATATATATATATATAAATA
P.t.  CCTTGGCTTCAGAAAAAA [AGCTTTTTTCTATATTTATAATTTTATATTTTCTATATATTTTATATTTTCCATATTATACCATTATATATATATAAATTAT
R.m.  CCTTGTTTTGGAAAAAAA [AAGTTTTATTTTATATATATATTTATAGATAAATTAAAAATTTAAATAGAAATT----------------------------
T.m.  CCTTGTTTTGAGAAAAAA [GGGTTTATTTTATATAATATATAATTATAGAAGATAAATATAAGAAAAAAATCCGAATAAAAT------------------
Z.m.  CCTTAGCTTAAAAAAAAA [AAACTTTTTTCTATATTTATAATTTTATATTTTCTATATATTTTAGATTTTCCATATTCTATTATACCATTATATATATATA
Z.n.  CCTTGGCTTCAGAAAAAA [AAATTTCTATATTTAGAATTTTTGATTTTCTATATATTTTAGATTTTCCATATTCTATTATACCATTATATATATATCAATT

[         200       210       220       230       240       250       260       270       280       290     ]
C.n.  ----------------------------------------------------------------------------------------------------
H.s.  AGTATATAGAGAAATAGAAATGGATTTCTTTTCTATATTGAAATATATTCCTATCCTTACGTATAGGATTGACATATAGGATTGACACCTGTAGAATAGGAG
P.o.  TATATATAATAAATTAAAATATATAATAAATT----------------------------------------------------------------------
P.t.  AAACTAAGAATT----------------------------------------------------------------------------------------
R.m.  ----------------------------------------------------------------------------------------------------
T.m.  ----------------------------------------------------------------------------------------------------
Z.m.  TATATAAATTAGAAACTAAGAATT----------------------------------------------------------------------------
Z.n.  ATAAACTAACAATT----------------------------------------------------------------------------------------
```

Fig. 1. Region of the chloroplast *trn*L intron sequenced in this study. Lower case letters indicate adjacent sequences. Conserved regions P, Q, R, and S are indicated. Bracketed strings of nucleotides (shown without alignment) were excluded from the analysis. Unbracketed bold face nucleotides comprise the 472 relatively conserved (alignable) sites used in phylogenetic analyses C.n. = *Cymodocea nodosa*; H.s. = *Halophila stipulacea*; P.o. = *Posidonia Oceanica*; P.t. = *Phyllospadix torreyi*; R.m = *Ruppia maritima*; T.m. = *Triglochin maritium*; Z.m. = *Zostera mariana*; Z.n. = *Zostera noltii*.

```
[       300         310  *******R*******330        340        350        360           370        380        390 ]
C.n.  ---] AAAAAA--GGATAGGTGCAGAGACTCAATGGAAGCTGTTCTAACGAATGAAGTTGTTTGCGTT [-----] GGTAGCAGGAATCTTTCTATCA---
H.s.  TTC] AAAAAA--GGATAGGTGCAGAGACTCAATGGAAGCTGTTCTAACGAATGGAGTTGGTTGCATT [GCGTC] GGTAATAGGGATCTTTCTATCG---
P.o.  ---] AAAAAAAGGGATAGGTGCAGAGACTCAATGGAAGCTGTTCTAACGAATGAAGTTGTTTGCGTT [-----] GGTAGCAGGAATCTTTCTATCA---
P.t.  ---] TTAAAA--GGATAGGTGCAGAGACTCAATGGAAGCTGTTCTAACGAATGAAGTTGTTTGCATT [-----] GGTAGCAGGAATATTTCTATCC---
R.m.  ---] TAAAAA-GGGATAGGTGCAGAGACTCGATGGAAGCTGTTCTAACGAATGAAGCTATTTGCATT [-----] GGTAGCAGGAATCTTTCTATCA---
T.m.  ---] AAAAAA--GGATAGGTGCAGAGACTCAATGGAAGCTGTTCTAACGAATAAAGTTGTTTGCGTT [-----] GGTAGCAGGAATCCTTCTATTAAAA
Z.m.  ---] TAAAAA-GGTATAGGTGCAGAGACTCAATGGAAGCTGT-CTAACGAATGAAGTTGTTTGCATT [T----] GGTAAGGGGAATAATTCTATCC---
Z.n.  ---] TTAAAA--GGATAGGTGAAGAGACTCAATGGAAGCTGT-CTAACGAATGAAGTTGTTTGCGTT [-----] GGTAGCAGGAATCTTTCTATCC---

[           400        410         420         430        440        450        460        470        480        490]
C.n.  --AAATTACAGAAAGGAAGGATGG [CCC-----] TATATACCTAAT-----ACATACGTATACATACTGACATATCAAATGATTAATCATAATCACGACC
H.s.  --AAATTGCCGAAAGGAAGAATGA [CCT-----] TATAGACCTAATATAATACATACGTATACATACTGACATATCAAACCATTCATCATAATCACAACC
P.o.  --AAATTACAGAAAGGAAGGATGA [CCCTATCC] TATATACCTAAT-----ACGTACGTATACATACTGACATATCAAACGATTAATCATAATCACGACC
P.t.  --AAATTACAGAAAGGAAAGATGA [CCC-----] TATATACCTAAT-----ACGTACATATACATACTGACATATCAAACGATTAATAATAATCACGACC
R.m.  --AAATTCCAGAAAGGAAGGATAA [ATC-----] TATATACCTAAT-----ACGTACGTATACATACTGACATATCAAACGATTCCTCATAATCACAACC
T.m.  TTAAATACACGAAAGAAAGGATAA [CCA-----] TATATACCCAAT-----ACATACGTATACATACTGACATATCAAACGATTAATCATAATCACGGCC
Z.m.  --GAATTACAGAAAGGAAAGATGA [CCT-----] TATATACCTAAT-----ACGTACATATACATACTGACATATTAAACGATCAATAATAATGACGACC
Z.n.  --TAATTACAGAAAGGAAAGATGA [CCC-----] TATATATCTCAT-----ACGTACATATACATACTGACATATTAAACGATTAATAATAATGACGACC

[             500         510        520        530         540         550         560        570        580    ]
C.n.  TTAATTTAATATATAATT [TATCTATAAATAAA--------] TTATATGAAAAATT [TACGGA---] GTTATT-ATGAATCCATGCTAACCCAAGTTG
H.s.  TGAATCTGGAATAGATAC [AAGAAATG--------------] TTCCATGAAAAATT [TCATGAATG] ATTGTGAATGAAGCCATATTCATCTCAGTTT
P.o.  TTAATCTAGAATTTATAT [ATAAATAAA-------------] TTCTATGAAAAATT [TACGGA---] GTTATT-ATGAATCCATGCCAACCCAAGTTG
P.t.  TTAATTTAGAATTTATAT [AAAATGGA--------------] TTCTATGAAAAATT [TACGGA---] CTTATT-GTGAATCCATGCCAATAAAAGTTG
R.m.  TGAATTTAATCTAGAATT [AAATTCTATTTAATTAAATAAA] TTCTGTGAAAAATT [GACGAACA-] GTTATT-ATGAATCCATG---TCCCAAGTTG
T.m.  TTAATCTAGAATTTATAT [AGAAATAAA-------------] TTCTACCTAAAATT [TTAGA----] ATTATT-GGAAATCCATGGCAATCCAAGTTG
Z.m.  TTAATTTAGAATTTATAT [AAAAAAAAA-------------] TTCTATGAAAATTT [TACGGA---] TTTCTT-GTGAATCCATGCCAATAAAAGTGG
Z.n.  TTAATCTAGAATTTATAT [AAAATAGA--------------] TTCTATGAAAAATT [TACGA----] TTTCTT-GTGAATCCATGCCGATGAAAGTGG
```

Fig. 1. (*Continued*)

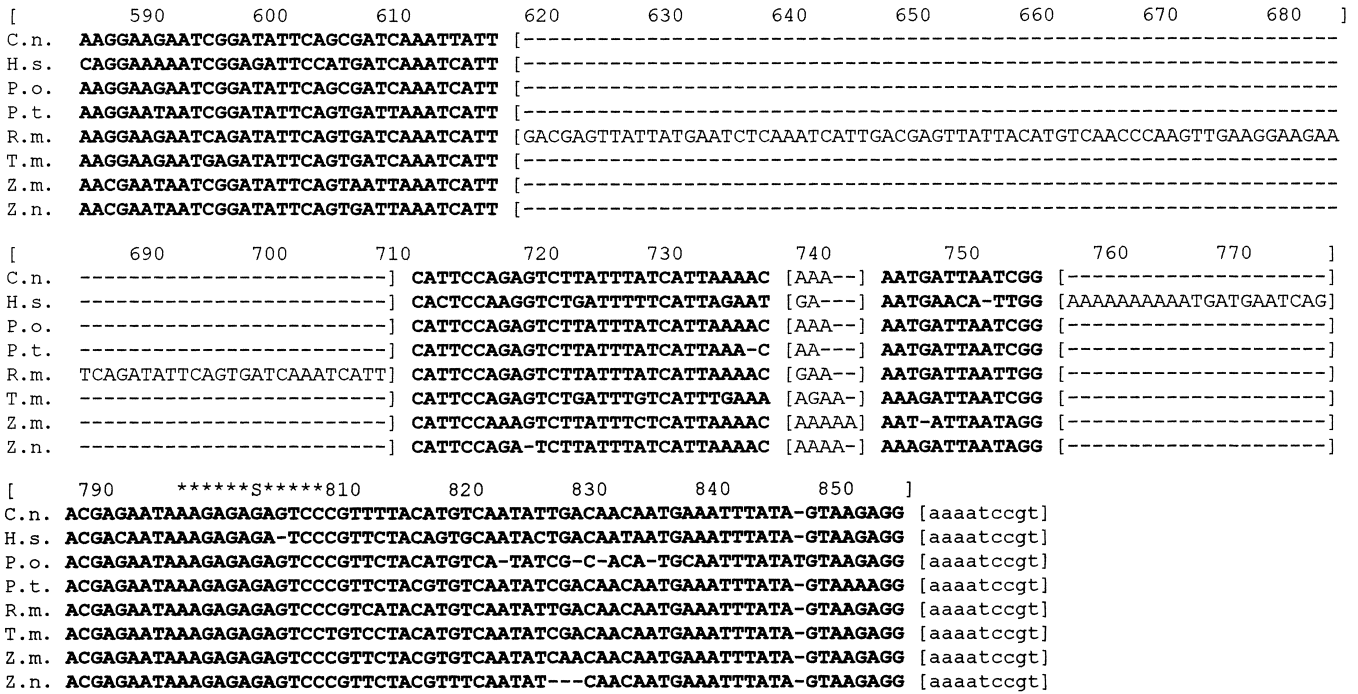

Fig. 1. (*Continued*)

The intron secondary structure contained four conserved regions (identified as P, Q, R and S) and nine stem regions, five of which ended in a loop (Kuhsel et al., 1990; Ferris et al., 1993). No variable sites occurred within the P and Q regions of the species studied, although two variable sites were present in the R region (*R. maritima* and *Z. noltii*) and one in the S region (*H. stipulacea*) (Fig. 1). Intron regions lacking secondary constraints contained large insertions/deletions (indels) and could not be aligned confidently. In particular, the region comprising the stem and loop 6 (between Q and R conserved regions) ranged from 90 bp in *R. maritima* to 220 bp in *H. stipulacea*. This large size difference is mainly attributable to a loop region where reliable alignment was possible only among closely related (i.e. co-familial) genera (Zosteraceae). Large size discrepancies also occurred in the region between the conserved R and S regions, where *R. maritima* and *H. stipulacea* possess 91 and 21 bp indels, respectively. *Posidonia oceanica* possessed a dinucleotide repeat in this variable region. This repeated sequence [$(TA)_6TTA(TA)_8TAAA$-$(TA)_4$] displays intra- and interpopulational polymorphism and has been utilized for population genetic analysis (Procaccini and Mazzella, 1998; Procaccini and Waycott, in press). Other minor gaps (indels of 1–15 bases) were scattered throughout the intron.

The region included in our phylogenetic analysis was 472 bp long (Fig. 1). Of the 472 nucleotide positions compared for the taxa examined, 125 sites (27%) were variable, 51 sites (10.8%) were informative and 74 sites (15.7%) were uninformative. For 1183 positions of *rbc*L compared with the same set of taxa, 314 sites (27%) were variable, 94 sites (8.0%) were informative and 220 sites (18.6%) were uninformative.

For *trn*L, pairwise divergence (uncorrected '*p*' distance) among taxa ranged from 4.2 (*Cymodocea* × *Posidonia*) to 16.7% (*Halophila* × *Zostera noltii*). For the same taxa, divergence was somewhat less for *rbc*L with '*p*' ranging from 1.4 (*Zostera noltii* × *Z. marina*) to 11.6% (*Halophila* × *Zostera noltii*). The range in divergence for the combined data was intermediate (Table 3). Skewness (g_1; an indicator of phylogenetic signal) was high for all datasets ranging from −0.936183 for *rbc*L, −0.968011 for *trn*L and −0.980419 for the combined data.

Trees and statistics resulting from our analyses are summarized in Figs. 2 and 3. All analyses resolved clades for Zosteraceae (*Phyllospadix*, *Zostera*) and *Zostera* (*Zostera marina*, *Z. noltii*) with high internal support (bootstrap 98–100%). A clade comprising *Ruppia* and *Cymodocea* was obtained in all analyses except unweighted parsimony of *trn*L sequences. Bootstrap support for this clade varied from 29 to 79%. Excluding the outgroup (*Halophila*), *Triglochin* (Juncaginaceae) occurred basal to the remaining seagrass genera in all analyses except unweighted parsimony of *trn*L sequences. In the combined analyses, bootstrap support for the basal position of *Triglochin* with respect to the seagrass genera reached 83% (Fig. 3(C)).

4. Discussion

Seagrasses possess similar morphological and physiological features that facilitate their survival in marine habitats (Les et al., 1993; Philbrick and Les, 1996). The possibility of convergent evolution of morphological characters in this group of plants has led to a number of different hypotheses concerning their origins, phylogenetic

Table 3
Uncorrected ('*p*') distances for three datasets used to examine phylogenetic relationships of seagrasses: *rbc*L coding region, *trn*L intron, combined data

	1	2	3	4	5	6	7	8
*rbc*L *coding region*								
1.*Cymodocea nodosa*	—							
2. *Halophila stipulacea*	0.09782	—						
3. *Posidonia oceanica*	0.02851	0.09552	—					
4. *Phyllospadix torreyi*	0.05972	0.11243	0.05917	—				
5. *Ruppia maritima*	0.04279	0.09721	0.04649	0.06847	—			
6. *Triglochin maritimum*	0.06562	0.10059	0.05325	0.07861	0.06932	—		
7. *Zostera marina*	0.06444	0.11496	0.06678	0.02705	0.07016	0.07777	—	
8. *Zostera noltii*	0.06540	0.11597	0.06512	0.02453	0.07019	0.07782	0.01439	—
*trn*L *intron*								
1. *Cymodocea nodosa*	—							
2. *Halophila stipulacea*	0.14095	—						
3. *Posidonia oceanica*	0.04149	0.12884	—					
4. *Phyllospadix torreyi*	0.07821	0.14768	0.05033	—				
5. *Ruppia maritima*	0.06552	0.13781	0.07030	0.09415	—			
6. *Triglochin maritimum*	0.09747	0.14985	0.06986	0.09537	0.10704	—		
7. *Zostera marina*	0.11125	0.16569	0.08104	0.04812	0.12260	0.12846	—	
8. *Zostera noltii*	0.10742	0.16672	0.07287	0.04397	0.12375	0.11842	0.04409	—
*Combined (rbc*L + *trnL) data*								
1. *Cymodocea nodosa*	—							
2. *Halophila stipulacea*	0.11103	—						
3. *Posidonia oceanica*	0.03245	0.10484	—					
4. *Phyllospadix torreyi*	0.06539	0.12234	0.05671	—				
5. *Ruppia maritima*	0.04968	0.10855	0.05313	0.07563	—			
6. *Triglochin maritimum*	0.07543	0.11441	0.05789	0.08335	0.07984	—		
7. *Zostera marina*	0.07862	0.12910	0.07073	0.03290	0.08475	0.09190	—	
8. *Zostera noltii*	0.07811	0.13007	0.06726	0.02993	0.08498	0.08909	0.02260	—

In pairwise comparisons of taxa, uncorrected sequence divergence of the *trn*L intron region compared to the *rbc*L coding region ranged from 0.85 : 1 (*Posidonia* × *Phyllospadix*) to 3.06 : 1 (*Z. marina* × *Z. noltii*). Uncorrected distances were essentially intermediate in pairwise comparisons of taxa for the combined data set.

relationships and evolution (den Hartog, 1970; Larkum and den Hartog, 1989; Cox and Humphries, 1993; ; Les and Haynes, 1995; Les et al., 1993, 1997; Philbrick and Les, 1996; Waycott and Les, 1996). Only one molecular dataset has been used previously to study broad seagrass relationships (Les et al., 1997). Since corroborating data are crucial for establishing the validity of phylogenetic hypotheses, the evaluation of different molecular data sets for studies of seagrass relationships is necessary. Chloroplast DNA $tRNA^{Leu}$ (UAA) intron sequences potentially offer one additional source of phylogenetically informative data to consider in questions regarding seagrass phylogeny.

The Group I intron internal to the tRNA leucine gene is conserved in terms of location, primary sequence and secondary structure (Kuhsel et al., 1990; Ferris et al., 1993). Size differences among members of the subclass Alismatidae that were examined reflect various insertion and deletion events that occurred during their extensive evolutionary history. Large indels were observed in loops associated with paired regions of the intron

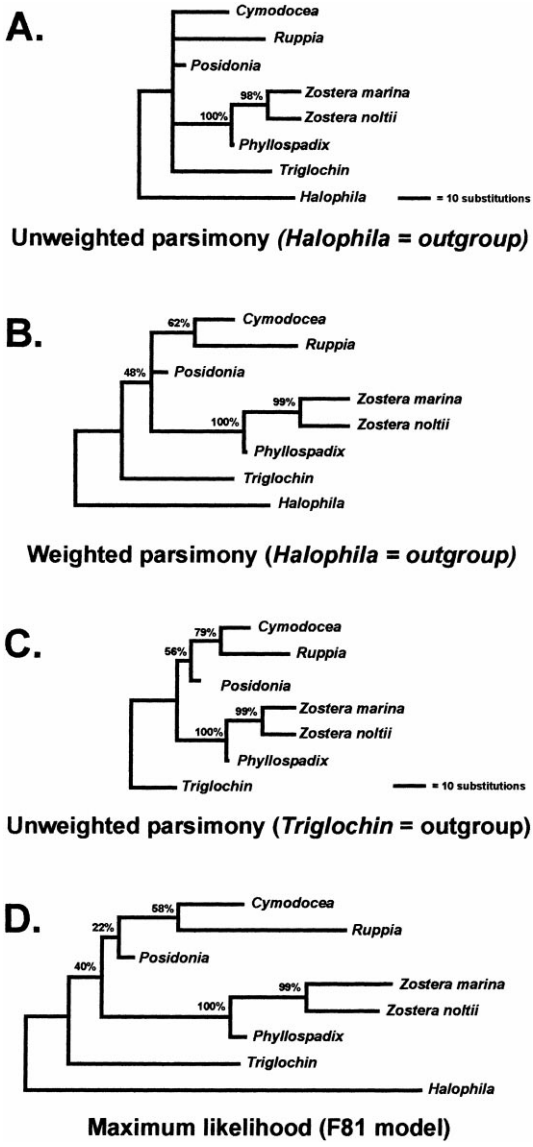

Fig. 2. Comparison of phylogenetic trees constructed for eight alismatid taxa using 472 conserved nucleotides from the chloroplast *trn*L intron. (A) Strict consensus of three minimum length trees (172 steps) obtained by maximum parsimony analysis unweighted for transition bias (CI = 0.84; $CI_{(EXC)}$ = 0.69; RI = 0.64). (B) Single minimum length tree (252 steps) resulting from weighted maximum parsimony analysis applying a 2 : 1 weight of transversions over transitions (CI = 0.85; $CI_{(EXC)}$ = 0.70; RI = 0.66). (C) Single minimum length tree (121 steps) resulting from unweighted maximum parsimony analysis excluding *Halophila* with *Triglochin* as the outgroup (CI = 0.89; $CI_{(EXC)}$ = 0.78; RI = 0.77. (D) Topology resulting from maximum likelihood analysis, F81 model (−ln = 1507.62998); the HKY model produced the same topology (−ln=1490.16086). Bootstrap percentages (shown) represent 500 replicates (A–D). Branch lengths are proportional to scale indicating 10 substitutions for unweighted analyses (A, C).

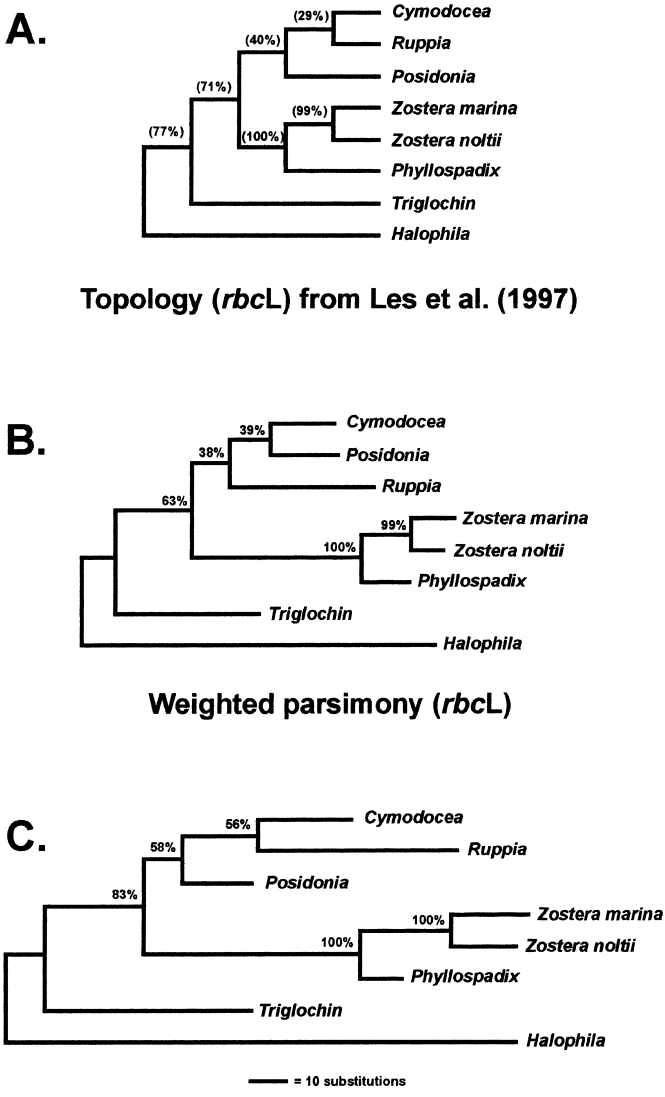

Fig. 3. Comparison of phylogenetic reconstructions for eight alismatid taxa using chloroplast *rbc*L coding and *trn*L intron sequences. (A) Topology based on maximum parsimony cladogram obtained in *rbc*L analysis of 79 taxa (adapted from Les et al., 1997). (B) Single minimum length tree (688 steps) resulting from weighted maximum parsimonny analysis (1183 *rbc*L nucleotides) of taxa analyzed in Fig. 1 (CI = 0.83; $CI_{(EXC)} = 0.70$; RI = 0.68). (C) Single minimum length tree (468 steps) resulting from maximum parsimony analysis of combined *trn*L and *rbc*L data (unweighted analysis shown; CI = 0.82; CI(EXC) =0.68; RI = 0.65). For combined data, the same topology and comparable nodal support were obtained in both weighted or unweighted parsimony analyses and maximum likelihood analyses (F81 or HKY models). Bootstrap percentages of B and C represent 500 replicates; bootstrap values in 'A' (in parentheses) represent comparable nodes on a larger cladogram of 79 taxa (Les et al., 1997).

(paired elements P1–P9), but a few differences were observed in four relatively conserved regions (P, Q, R, S; Kuhsel et al., 1990). We observed that much of the variation in intron size was due to large indels in the P6 and P8 loops. Intra-specific variability of these regions has recently been exploited to study disjunct *Posidonia oceanica* populations of the Western Mediterranean basin, where the microsatellite locus (*Poc-trn*) occurs with different alleles varying in size by 14 bases (Procaccini and Mazzella, 1998).

By avoiding hypervariable regions as indicated by the secondary structure of the *trn*L intron, we were able to select conserved regions (mainly paired elements) that were more amenable to higher levels of phylogenetic analysis such as intergeneric seagrass relationships. Unambiguous alignment of these regions was possible among even the more distantly related taxa included in our analysis. A much larger portion of the *trn*L intron sequence could have been used had we restricted our analysis only to closely related taxa as advocated by Gielly and Taberlet (1994).

Although limited by its taxonomic representation, our analysis allows us to assess the degree of phylogenetic resolution attainable through use of conserved regions of the *trn*L intron. This was achieved by comparison with phylogenetic trees constructed using conserved DNA sequences from the coding regions of the chloroplast *rbc*L gene. Since both types of sequences are located in the chloroplast, they necessarily share the same evolutionary history. Therefore, in the absence of systematic errors (Swofford et al., 1996), the sequences should yield congruent phylogenetic reconstructions. The *rbc*L and *trn*L datasets are, in fact, highly congruent ($I_{MF} = 0.01$). Nevertheless, for this particular group of taxa, the intron evolves from 1.4 to 3 times faster than the coding sequence, in agreement with the average values of 2.71 indicated by Gielly and Taberlet (1994), where the *trn*L intron was found to evolve at a rate intermediately between coding sequences and intergenic spacers.

Unweighted parsimony analysis of the *trn*L intron data alone results in weak resolution of distantly related taxa, but excellent resolution among closely related taxa such as members of the family Zosteraceae (Fig. 2(A)). Enhanced phylogenetic resolution (Fig. 2(B)) is obtained by downweighting transitions in a manner similar to that used to weight third position sites in coding regions of cpDNA (Les et al., 1997). As parsimony methods are known for poor performance in instances of long branch 'attraction' (Felsenstein, 1978), several other analyses were performed. First, we removed the *Halophila* sequence which was the most divergent sequence indicated by pairwise comparisons of uncorrected '*p*' distances (Fig. 3). This, in effect, greatly reduced the potential for long branch attraction to other sequences. Repeating the unweighted parsimony analysis (now using *Triglochin* as the outgroup; Fig. 2(C)), we recovered the same topology among taxa as a previous analysis of *rbc*L sequences for a larger set of taxa (Fig. 3(A)). We also evaluated the complete dataset (restoring *Halophila* as the outgroup) using a maximum likelihood algorithm which is not as susceptible to long branch artifacts. Evaluating models either with minimal assumptions (F81) or different rate assumptions (HKY), the likelihood approach recovered the same topology (Fig. 2(D)) as the *rbc*L analysis (Fig. 3(A)) even with the distant outgroup included. These experiments demonstrate the problem of long branch attraction in data sets with relatively incomplete taxonomic sampling, but illustrate how the problem can be resolved by removing the long branch or by using maximum likelihood approaches for reconstructing the phylogenetic relationships.

Taxon 'density' (taxon sampling) also influences the results of phylogenetic analyses. Analyzing *rbc*L data for the subset of eight taxa (Fig. 3(B)) yielded a slightly different topology than analyses including 79 taxa (Les et al., 1997; Fig. 3(A)). Here, the same clades were resolved, but the positions of *Ruppia* and *Posidonia* were reversed.

Analysis of combined *rbc*L and *trn*L data resulted in a single most parsimonious tree with the same topology as the *rbc*L analysis of 79 taxa. Several nodes in the cladogram had excellent support (83–100%) and two nodes received good support (56–58%). Although the latter values are not exceptionally high, they are 45–93% higher than the values obtained using *rbc*L data alone. The combined data also had a higher value of skewness (g_1) indicating an increase in phylogenetic signal over the individual datasets. These results enforce the utility of combined data analysis not only to corroborate, but to provide greater support for existing phylogenetic hypotheses.

Our analyses indicate that these combined data strengthen specific hypotheses of seagrass relationships that have been established previously. The freshwater and bisexual Juncaginaceae appear as a basal group with respect to the marine families Cymodoceaceae, Posidoniaceae, Ruppiaceae and Zosteraceae, in agreement with the hypothesis that seagrasses evolved from fresh water ancestors (Les et al., 1997; but see also Larkum and den Hartog, 1989). Moreover, our results support the hypothesis that bisexuality occurs as the primitive floral condition of seagrasses in the Cymodoceaceae 'complex' which comprises Cymodoceaceae, Posidoniaceae and Ruppiaceae (Les et al., 1997). This study provides additional support to include the hermaphroditic Ruppiaceae near the base of the Cymodoceaceae 'complex' clade which includes the hermaphroditic Posidoniaceae and dioecious seagrasses (e.g. Cymodoceaceae). Finally, the distinction of the Cymodoceaceae complex and Zosteraceae is confirmed by the high level of molecular divergence observed to occur between these groups.

To summarize, we have shown that conserved regions of the cp *trn*L intron can be used effectively to study intergeneric phylogenetic relationships among seagrasses and other members of the monocot subclass Alismatidae. In itself, the *trn*L intron can resolve even distant relationships when suitable methods of analysis are applied. Since *trn*L and *rbc*L sequences are highly congruent and share the same evolutionary history, their combination for phylogenetic studies is warranted.

Further sequencing of the *trn*L region is encouraged for other alismatid taxa and should provide additional support for the emerging hypothesis of seagrass relationships initiated by *rbc*L data analysis. Since variable and conserved portions of the *trn* L intron can be readily partitioned, this versatile region provides a range of genetic markers for studying relatively distantly related taxa, closely related taxa, or even intraspecific relationships. The small size of the intron, existence of variably divergent regions and availability of universal primers for PCR amplification and sequencing provide an ideal combination of features for application in evolutionary studies of seagrasses.

Acknowledgements

We thank D.L. Robertson and G.J. Smith for continuous support during the first part of the work and J.L. Olsen for useful comments on the manuscript.

References

Chase, M.W., Soltis, D.E., Olmstead, R.G., Morgan, D., Les, D.H., Mishler, B.D., Duvall, M.R., Price, R.A., Hills, H.G., Qiu, Y.-L., Kron, K.A., Rettig, J.H., Conti, E., Palmer, J.D., Manhart, J.R., Sytsma, K.J., Michaels, H.J., Kress, W.J., Karol, K.G., Clark, W.D., Hedrén, M., Gaut, B.S., Jansen, R.K., Kim, K.-J., Wimpee, C.F., Smith, J.F., Furnier, G.R., Strauss, S.H., Xiang, Q.-Y., Plunkett, G.M., Soltis, P.S., Swensen, S.M., Williams, S.E., Gadek, P.A., Quinn, C.J., Eguiarte, L.E., Golenberg, E., Learn Jr, G.H., Graham, S.W., Barrett, S.C.H., Dayanandan, S., Albert, V.A., 1993. Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. Ann. Miss. Bot. Gard. 80, 528–580.

Cox, P.A., Humphries, C.J., 1993. Hydrophilous pollination and breeding system evolution in seagrasses: A phylogenetic approach to the evolutionary ecology of the cymodoceaceae. Bot. J. Linn. Soc. 113, 217–226.

Cronquist, A., 1988. The Evolution and Classification of Flowering Plants, New York Botanical Garden, New York, 555pp.

Den Hartog, C., 1970. The Seagrasses of the World, North Holland, Amsterdam, 275pp.

Felsenstein, J., 1978. Cases in which parsimony and compatibility methods will be positively misleading. Syst. Zool. 27, 401–410.

Ferris, C., Oliver, R.P., Davy, A.J., Hewitt, G.M., 1993. Native oak chloroplast reveals an ancient divide across Europe. Mol. Ecol. 2, 337–344.

Giannasi, D.E., Zurawski, G., Learn, G.H., Clegg, M.T., 1992. Evolutionary relationships of the caryophillidae based on comparative *rbcL* sequences. Syst. Bot. 17, 1–5.

Gielly, L., Taberlet, P., 1994. The use of chloroplast DNA to resolve plant phylogenies: Noncoding versus *rbc*L sequences. Mol. Biol. Evol. 11(5), 769–777.

Heery, D.M., Gannon, F., Powel, R., 1990. A simple method for subcloning DNA fragments from gel slices. TIG 6, 171.

Higgins, D.G., Bleasby, A.J., Fuchs, R., 1992. CLUSTAL V: Improved software for multiple sequence alignment. Comput. Appl. Biosci. 8, 189–191.

Kuo, J., McComb, A.J., 1989. Seagrass taxonomy, structure and development, In: Larkum, A.W.D., McComb, A.J., Sheperd S.A., (Ed.), Biology of Seagrasses, Elsevier, Amsterdam, pp. 5–73.

Kuhsel, M.G., Strickland, R., Palmer, J.D., 1990. An ancient group I intron shared by Eubacteria and chloroplasts. Science 250, 1570–1573.

Larkum, A.W.D., den Hartog, C., 1989. Evolution and biogeography of seagrasses, In: Larkum, A.W.D., McComb, A.J., Sheperd, S.A., (Ed.), Biology of Seagrasses, Elsevier, Amsterdam, pp. 112–156.

Les, D. H., Haynes, R. R., 1995. Systematics of subclass Alismatidae: A synthesis of approaches, In: Rudall, P.J., Cribb, P.J., Cutler, D.F., Humphries, C.J., (Ed.), Monocotyledons: Systematics and Evolution, Royal Botanic Gardens, Kew, UK, pp. 353–377.

Les, D.H., Garvin, D.K., Wimpee, C.F., 1993. Phylogenetic studies in the monocot subclass Alismatidae: Evidence for a reappraisal of the aquatic order Najadales. Mol. Phylogenet. Evol. 2, 304–314.

Les, D.H., Cleland, M.A., Waycott, M., 1997. Phylogenetic studies in Alismatidae, II: Evolution of marine angiosperms ('seagrasses') and hydrophily. Syst. Bot. 22, 443–463.

Philbrick, C.T., Les, D.H., 1996. Evolution of aquatic angiosperm reproductive systems. BioScience 46(11), 813–826.

Procaccini, G., Alberte, R.S., Mazzella, L., 1996. Genetic structure of the seagrass *Posidonia oceanica* (L.) Delile in the Western Mediterranean: Ecological implications. Mar. Ecol. Prog. Ser. 140, 153–160.

Procaccini, G., Mazzella, L., 1998. Genetic structure and gene flow in *Posidonia oceanica* (L.) Delile populations from the West Mediterranean basin: A microsatellite analysis. Mar. Ecol. Prog. Ser. 169, 133–141.

Procaccini, G., Waycott, M., in press. Microsatellite loci identified in the seagrass *Posidonia oceanica* (L.) Delile. J. of Hered.

Soltis, D.E., Soltis, P.S., Milligan B.G., 1992. Intraspecific chloroplast DNA variation: Systematic and phylogenetic implications, In: Soltis, D.E., Soltis, P.S., (Eds.), Sinauer, New York, pp. 117–150.

Swofford, D.L., 1991. When are phylogeny estimates from molecular and morphological data incongruent? In: Miyamoto, M.M., Craycraft, J., (Eds.), Phylogenetic Analysis of DNA Sequences, Oxford University Press, New York, pp. 295–333.

PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1.1., Illinois Natural History Survey, Champaign, IL.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic Iinference, In: Hillis, D.M., Moritz, C., Mable, B.K., (Eds.), Molecular Systematics, 2nd edn., Sinauer, Sunderland, MA, pp. 407–514.

Taberlet, P., Gielly, L., Pautou, G., Bouvet, J., 1991. University primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17, 1105–1109.

Yong-Pyo, H., Hipkins, V.D., Strauss, S.H., 1993. Chloroplast DNA diversity among trees, populations and species in the California closed-cone Pines (*Pinus radiata, Pinus maricata* and *Pinus attenuata*). iChloroplast DNA diversity among trees, populations and species in the California closed-cone pines (*Pinus radiata, Pinus maricata* and *Pinus attenuata*). Genetics 135, 1187–1196.

Wolfe, K.H., Li, W.-H., Sharp, P.M., 1987. Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. Proc. Natl. Acad. Sci. U.S.A. 84, 9054–9058.

Waycott, M., Les, D.H., 1996. An integrated approach to the evolutionary study of seagrasses, In: Kuo, J, Phillips, R.C., Walker, D.I., Kirkman, H., (Eds.), Seagrass Biology: Proc. Int. Workshop, Rottnest Island, Western Australia, pp. 71–78.

Zurawski, G., Clegg, M.T., 1987. Evolution of higher-plant chloroplast DNA-encoded genes: Implications for structure–function and phylogenetic studies. Ann. Rev. Plant Phys. 38, 391–418.